Bringing the power of genetic research to an office near you

April 10, 2016

Science on the Hill: Bringing the power of genetic research to an office near you

by Patrick Chain

Most of us have gone to the doctor ourselves or taken our children with a sore throat and sinus congestion, only to find our physician couldn't readily tell whether we had a cold virus or a bacterial infection. Just in case, we might have walked away from the appointment with a possibly unnecessary prescription for antibiotics.

Now imagine that a nurse could swipe your saliva and run a quick genetic test for bacteria. If it comes back negative, this time you walk out with just a script for decongestant and orders to get some rest instead of buying an unnecessary prescription and contributing to the antibiotic resistance crisis.

That's just one example of the benefits of rapid genetic screening on a personal level. On a grander scale, the ability to quickly analyze genetic data stands to revolutionize research into everything from the mutations causing various cancers to the "Second You," your microbiome, or the bacteria living inside you. Genomics can also revolutionize our understanding of a range of diseases—Alzheimer's, irritable bowel syndrome, Crohn's disease, for instance—as well as how to grow algae to best produce oil to make gasoline. In medicine, genetic screening can tell hospital staff what pathogens inhabit the hospital environment. In environmental research, it can clarify how communities of microorganisms fix carbon from the atmosphere and how their populations adapt to less rain and hotter summers.

Genomics—the genetic mapping and DNA sequencing of sets of genes or the complete genomes of organisms, along with related genome analysis and database work—

are emerging as one of the transformative sciences of the 21st century. Partly that's resulting from the rapid spread of so-called next-generation sequencing instruments, which have become accessible to the average biologist and, eventually, to the physician. Gene sequencing has become much more democratized over the last few years.

Decreasing costs for sequencing instruments is driving their spread to new users, making them available to the common scientist. Today you'll find sequencers not only in most universities and other large research institutions, but also in hospitals, individual clinics and the small labs of individual researchers. Genomics has become the cornerstone of all biological research, which almost always involves sequencing all

the genes (the genome) of the organism under study or the many species forming a community to see what's going on. So everyone wants their own capability in-house.

All this easily and rapidly generated data has caused a new bottleneck, as the ability to analyze the data—and it's very big Big Data—is swamping genomics. Bioinformatics tools use computers to pull together, classify, store, process, and analyze molecular genetic and genomic data. Unfortunately, the current tools are not entirely user-friendly or accessible to most biological researchers, who have more expertise in biology than in crunching data.

Seeing a need that the unique expertise at Los Alamos National Laboratory could fill, a team in the Biosecurity and Public Health group, collaborating with the Naval Medical Research Center, has developed a new computational and web-based tool called EDGE Bioinformatics to fulfill the promise of democratizing genomics.

Funded by the Department of Defense's Defense Threat Reduction Agency, the work comes out of the Lab's decades of research in genetics and life sciences. Long interested in the link between radiation and genetic mutations, the U.S. Department of Energy (DOE) and the National Institutes of Health received federal funding in 1998 to begin the Human Genome Project to sequence, or map, the genome of the species Homo sapiens—us. Los Alamos was a key player, contributing its expertise in life sciences, particularly genetics, and its world-class computing resources to the task of unraveling the human genetic code. By June 2003, the map was mostly complete. Since then, the Lab has applied its expertise to a range of related genetic research, from illuminating the causes of cancer to perfecting algae for biofuel production.

For the Los Alamos EDGE team, it was a natural step from this background to creating a handy, easy-to use, web-based computer program with a wide assortment of integrated and pioneering bioinformatics tools. EDGE includes several pre-configured workflows to analyze sequencing data, identify genomes, and create reports and graphics based on the data. Using EDGE, with a few mouse clicks a novice in bioinformatics can create sophisticated analyses of a sample in minutes instead of days or weeks.

This bioinformatics platform was designed as an initial attempt at empowering the development of genomics expertise—that's what EDGE stands for. EDGE has already helped streamline data analysis for groups in multiple countries worldwide as well as within several government laboratories in the United States. Because the program is "open source," anyone can use it or even modify it to suit their needs and bring the power of Big Data Analysis to even the smallest research lab—or doctor's office.

Genomics researcher Patrick Chain is the EDGE team leader in the Biosecurity and Public Health group at Los Alamos National Laboratory. With a background in microbial ecology, evolution, genomics and bioinformatics, Chain has spent the past 20 years using genomics to study various microbial systems. He currently leads a team of researchers whose charge is to devise novel methods, algorithms and strategies for the biological interpretation of massively parallel sequencing data. Read more about how EDGE is making genetic sequencing analysis widely available to research scientists, geneticists, health organizations of all sizes, and the medical community at large. Read more about the Human Genome Project at the National Human Genome Research Institute.

This article first appeared in the Santa Fe New Mexican.